Patent Idea Details

FILE COPY

GENERAL INFORMATION

      Title: URL-Based TTS

      ID:

  Patent No.: —, —

      URL: [Application No. —]

  Inventors: <u>Lewis Dodrill (ddodrill)</u>, <u>Ryan Danner (rdanner)</u>, and <u>Steven Martin (sjm)</u>
      More details on these inventors <u>listed below</u>.

    Date Entered: 22-Feb-2000

    Date Modified: 22-Feb-2000

  Date Filed: —

  Date Issued: —

Background: There currently exist a variety of software and hardware solutions that convert text information into audio in the form of human speech. This "text-to-speech" conversion serves a variety of useful purposes such as providing textual information to the visually impaired or to non-visual devices such as telephones. There are, however, inherent shortcomings in the existing technology that can make it difficult to integrate into applications or exhibit undesirably low levels of performance.

The shortcomings of existing text-to-speech implementations are consistently exposed. For example, the most straightforward conversion currently being implemented is to send a body of text to a text-to-speech converter and wait for an audio file to be generated. This, unfortunately, means waiting for the full conversion process to complete. The common solution to this problem is to wait for a certain amount of the conversion process to complete and then begin playing the generated audio as it's converted. While faster, it makes the assumption that the conversion process is fast enough to "keep up" with the audio being played and, furthermore, makes the assumption that the software component receiving the audio information has direct access to the audio device (pc soundcard, telephone resource, etc) to play the audio. The invention outlined herein provides a more elegant solution that takes advantage of the HTTP standard by providing URL access to a text-to-speech resource in such a way that both simplifies software integration and application performance. Furthermore, by HTTP-enabling existing TTS software, performance and ease of use is improved without having to re-write the complicated "engines" that convert the text and generate the audio.

1 of 4

Summary: The idea of this invention is to provide a web-server-based implementation of text-to-speech conversion via HTTP-standard URL's. Existing "text-to-wav" or "text-to-media" software that exists today could be used as the basis of such a server. The web server must contain a CGI application or server module with the three following capacities:

1. The ability to convert a body of text received via http post to the main URL of the application into a series of smaller URLs each containing part of the text to be converted.

2. A text-to-media capability either natively or through invocation of existing software residing on the web server.

3. A mechanism to convert a URL generated by item number 1 utilizing the capability described by item number 2.

In order for an application to access the server described above, it would first POST the text to be converted to the base URL of the server. The process would proceed as follows:

The text "This is a sample body of text. It contains multiple sentences, and a variety of punctuation; all of these are elements of standard text that are interpreted by text-to-speech converters." would be sent via a FORM POST to http://ttsserver/main.cgi . The tts server would then return the following list or urls "http://ttsserver/main.cgi?This+is+a+sample+body+of+text%p++It+contains, http://ttsserver/main.cgi?+multiple+sentences%c+and+a+variety+of+punctuation%c, http://ttsserver/main.cgi?+all+of+these+are+elements+of+standard+text+that+are, http://ttsserver/main.cgi?+interpreted+by+text%dto%dspeech+converters.". The application would then use HTTP GET requests to the returned URLs in sequence. The data returned from the GET would be standard audio information such as WAV data. If desired, the application could GET the next URL in the sequence while playing the data returned by the preceding GET. If the TTS process needs to be interrupted, the application may stop requesting URLs at any point in the list.

Advantages: This invention improves many facets of text-to-speech services. By breaking up the text into multiple URLs from which the converted text media is retrieved, implementation of text-to-speech is greatly simplified. Furthermore, any standard audio-enabled web client can retrieve the converted audio natively. Another advantage is that the http standard allows for each request can be divided among multiple servers to enhance tts-resource availability and load balancing.

Method of Detecting Use By Other Companies: A protocol analyzer can be used to look for large URLs returning audio data. Also, web servers co-resident with existing text-to-speech software could be examined to check for interaction between the web server and the TTS software.

Previous Public Use: ---

First Written Record Date: ---

2 of 4